What does investors' online divergence of opinion tell us about stock returns and trading volume?

Article (Accepted Version)

What Does Investors' Online Divergence of Opinion Tell Us About Stock Returns and Trading Volume?☆

Alya AL-NASSERI[a] and Faek MENLA ALI[b*]

[a] *Salalah College of Technology, Sultanate of Oman*
[b] *School of Business Management and Economics, University of Sussex, Brighton BN1 9SL, UK*

January 2018

Abstract

We analyse 289,443 online tweets from StockTwits and construct a divergence of opinion (disagreement) indicator for investigating the impact of disagreement on stock returns and trading volume. We find that the impact of disagreement on returns is asymmetric; it is negative (positive) during bull (bear) market periods. We also find that higher online disagreement increases trading volume; this effect is detected irrespective of whether the market is bullish or bearish. Moreover, portfolio strategies that are designed on the basis of our disagreement indicator are shown to generate abnormal profits. Overall, our results confirm the important role of belief dispersion in financial markets.

Keywords: Disagreement, Online tweets, Stock returns, Trading volume

JEL Classification: C21, G02, G14

☆ We would like to thank the Editor-in-Chief (Professor Anders Gustafsson), the Associate Editor (Professor Karthik Krishnan) and the anonymous referee for the valuable comments and suggestions on an earlier draft of this manuscript. We are also grateful for their useful comments and suggestions to Professor Guglielmo Maria Caporale and participants of the Portsmouth-Fordham Conference on Banking and Finance, 24-25 September 2016, the University of Portsmouth.

* Corresponding author (Faek Menla Ali) at: 207 Jubilee Building, School of Business Management and Economics, University of Sussex, Brighton BN1 9SL, UK. Tel: +44(0)1273 872969. *Email address*: F.Menla-Ali@sussex.ac.uk (F. Menla Ali).

1. Introduction

The impact of heterogeneity in investors' beliefs or opinions on stock market behaviour has attracted considerable attention in recent years. However, the available theoretical and empirical evidence is inconclusive. On the one hand, in the presence of short-sale constraints, lower expected returns are associated with divergence of opinion (Miller, 1977; Diether, Malloy, & Scherbina, 2002; Chen, Hong, & Stein, 2002; Berkman, Dimitrov, Jain, Koch, & Tice, 2009; Yu, 2011; among others). On the other hand, several studies argue that differences in opinions can lead to higher risk premia (e.g., Varian, 1985; Garfinkel and Sokobin, 2006; David, 2008). Disagreement has also been linked to the trading volume in asset markets. While the "no-trade theorem" of Milgrom and Stokey (1982) finds that it causes revisions of beliefs without affecting volume, subsequent studies that use different proxies for disagreement conclude that it increases the trading volume (see, e.g., Harris & Raviv, 1993; Kandel & Pearson, 1995; Bamber, Barron, & Stober, 1997; Banerjee & Kremer, 2010; Carlin, Longstaff, & Matoba, 2014; among others).

The present study aims to shed new light on the effects of disagreement on stock market behaviour by addressing the following specific questions: Does divergence of opinion among stock-related tweets affect stock returns and trading volume? Are returns and volume effects of divergence of opinion asymmetric? Does the divergence of opinion among tweets help to predict stock returns? Since the rise of social media platforms in recent years, a rapidly growing body of research has examined whether sentiment and disagreement indicators that are constructed from microblogging posts are associated with stock market features (for a comprehensive overview, see Bukovina, 2016).[1] For instance, Antweiler and Frank (2004) construct a disagreement indicator from Internet message boards and show that disagreement among messages increases trading volume. Zhang, Fuehres, and Gloor (2011) find that the percentage of emotional tweet posts is negatively correlated with various US stock market indices. Bollen, Mao, and Zeng (2011) show that the accuracy of Dow Jones Industrial Average (thereafter DJIA) predictions is significantly improved when certain public mood dimensions from Twitter are included. Sprenger, Tumasjan, Sandner, and Welpe (2014) use stock-related messages (from the so-called StockTwits) and find linkages between tweet sentiment and stock returns; message volume and trading volume; and disagreement and volatility. Giannini, Irvine, and Shu (2015) also construct a disagreement measure from Twitter posts and show that both divergence and

[1] There is also a large body of literature in which the impact of indicators that are extracted from news articles on stock market features is examined (see, for example, Tetlock, 2007; Fang & Peress, 2009; Yuan, 2015; among others).

convergence of opinion generate abnormal trading volume at the time of and after earnings announcements.[2]

This paper contributes to this area of the literature by providing further evidence for the role of online divergence of opinion in the prediction of stock returns and trading volume. For example, Antweiler and Frank (2004) extract their disagreement indicator from Internet message boards; however, as Sprenger et al. (2014) note, such boards require users to actively access the forum for up-to-date developments on a particular stock and the information becomes outdated in the absence of new posts. More recent studies analyse instead the impact of online tweets that are posted on Twitter. However, most of them employ a randomised subsample of all posted tweets and/or focus on stock market feature effects of microblog-extracted indicators (i.e., sentiment, attention, and mood) other than disagreement (e.g., Zhang et al., 2011; Bollen et al., 2011; Mao et al., 2011; Zhang, Li, Shen, & Teglio, 2016; among others). By contrast, we analyse 289,443 online tweets that are directly related to stock features (e.g., posted on the microblogging website StockTwits) and construct a disagreement or divergence of opinion indicator among these tweets to examine its effect on the returns and trading volume of stocks. Specifically, we use different classification algorithms from computational linguistics to classify the collected messages into three distinct classes M^c, where c $\in$ {Buy, Hold, Sell}. Then, a disagreement indicator among messages is constructed and used in the empirical analysis. The data set includes the 30 DJIA stocks over the period from April 4, 2012 to April 5, 2013. These stocks are highly liquid and characterised by high market capitalisation and institutional ownership, and their short-sale transactions represent a high percentage of their daily volume. That is, they are likely to be free from short-sale constraints and other market frictions and, therefore, particularly suitable for analysing the impact of disagreement on stock returns and trading volume.[3] Moreover, they generate a great buzz and are discussed very frequently on StockTwits; hence, extracting a disagreement indicator from the collected messages is

[2] Microblogging messages have also been analysed to establish whether they are associated with macroeconomic indicators (see, e.g., Bokányi, Lábszki, & Vattay, 2017). There is an ongoing stream of research that examines whether Internet search data (e.g., Google queries) operate as economic indicators - see, for example, McLaren and Shanbhogue (2011) for predicting housing and labour markets, D'Amuri and Marcucci (2017) for predicting the unemployment rate, Saxa (2014) for predicting mortgage lending, and Choi and Varian (2012) for predicting automobile sales, to name a few.

[3] Indeed, the general conclusion of the literature on short-sale activity is that short-sale constraints are more binding for stocks with low institutional ownership and low market capitalisation (see, e.g., D'Avolio, 2002; Diether, Lee, & Werner, 2009; among others).

particularly appropriate since they may contain valuable information about investors' divergence of opinion and sentiment.[4]

To the best of our knowledge, the only studies to date to have used a disagreement indicator that was constructed from StockTwits messages are those by Sprenger et al. (2014) and Giannini et al. (2015);[5] however, this paper differs from them in various crucial ways. Specifically, Sprenger et al. (2014) carry out the analysis for the companies that are listed on the S&P 100 index over the period from January 1, 2010 to June 30, 2010, whilst the present study examines a longer sample period using a different empirical approach and focuses on the 30 highly liquid stocks in the US that are most frequently discussed in online stock forums. Moreover, our disagreement indicator is more informative than that constructed by Giannini et al. (2015) since it (i) excludes the neutral (hold) and re-tweet messages, which may contain a certain amount of noise and, hence, distort the indicator, and (ii) reflects the true opinions of investors who are engaged in the decision-making process, rather than the impact that each post has through its followers (which may not reflect the actual opinions of the platform's participants).

Our findings are as follows: The effect of our online disagreement indicator on stock returns appears to be insignificant, which is consistent with the findings of Antweiler and Frank (2004). However, it does affect trading volumes, as was also found by Harris and Raviv (1993), Kandel and Pearson (1995), Antweiler and Frank (2004), Sprenger et al. (2014) and Carlin et al. (2014), among others.

Then, we extend the analysis to distinguish between (possibly asymmetric) disagreement effects on returns and volume in bull and bear markets. Although the behavioural finance literature has widely debated whether investors behave differently in different states of the economy or the market (e.g., Lee, Jiang, & Indro, 2002; Verma & Verma, 2007; Chung, Hung, & Yeh, 2012; among others), the existing empirical studies on disagreement effects on stock market features have only considered linear dependence (e.g., Antweiler & Frank, 2004; Sprenger et al., 2014; among others). We find that returns respond negatively (positively) to disagreement during bull (bear) periods. Moreover, a positive disagreement effect on volume is detected regardless of market conditions (i.e., bull vs. bear periods). To the best of our knowledge, our paper is the first to explore the asymmetric

[4] The recently emerged StockTwits forum has various distinct features, such as a high volume of message posts, messages are posted in real time, and an efficient diffusion mechanism of information and opinions among investors.

[5] Note that Sprenger and Welpe (2011) use such data in a different context, i.e., to analyse whether S&P 500 stock prices are associated with different company-specific news events that are published on Twitter (e.g., corporate governance or legal issues).

disagreement effects on returns and volume. It shows that, unlike disagreement effects on volume, which are found to be symmetric, such effects on returns are asymmetric.

Finally, we find that abnormal profits can be obtained when portfolio strategies are designed according to our disagreement indicator. The returns of portfolios of low to medium disagreement are higher than those of other portfolios, and this difference is highly significant in the case of stocks with relatively lower trading volume, which is consistent with the evidence that was reported by Sadka and Scherbina (2007). Previous empirical studies investigate profitable predictability in the cross-section of stock returns based on a wide range of measures for the divergence of opinion (for example, higher trading volume in Lee and Swaminathan (2000), breadth of mutual fund ownership in Chen et al. (2002), dispersion in analysts' earnings forecasts in Diether et al. (2002), Doukas, Kim, and Pantzalis (2006), Verardo (2009), Yu (2011), and Banerjee (2011) among others, and unexpected trading volume in Chen, Qin, and Zhu (2015), among others). This paper is the first to provide evidence on this issue using a disagreement indicator that is constructed from online tweets.

Our findings have important implications for practitioners. In particular, portfolio strategy design should take into account the asymmetric effects of disagreement in bull vis-à-vis bear markets and for stocks with different trading volumes.

The paper is organised as follows: Section 2 presents the data, the classification methods that were employed, and the measurement of the online divergence of opinion. Section 3 outlines the empirical methodology. Section 4 discusses the empirical results and some robustness checks. Section 5 examines the role of the online divergence of the opinion indicator in predicting the cross-section of stock returns. Finally, Section 6 presents the conclusions of the paper.

2. Data Description, Tweet Classification and Divergence of Opinion Indicator

2.1 StockTwits Data

In this study, we construct a divergence of opinion indicator from StockTwits data and analyse its effect on stock returns and trading volume. More specifically, one year of StockTwits data on the companies listed on the DJIA index are downloaded from the Application Programming Interface (API) website for the period April 4, 2012 – April 5, 2013, which consists of 251 days.[6] Over 3.5 million stock microblog posts were initially

[6] To manage the high volume of tweet posts, we focus on a one-year period, which is still longer than the six-month period

obtained from the API over the sample period. These posts were pre-processed, and posts without any ticker, those with more than one ticker, those that were re-tweeted and those that were not related to companies that are listed on the DJIA index were removed, thereby leaving 289,443 valid posts that contained the dollar-tagged ticker symbol of the 30 stock tickers of the DJIA index.

In addition, consistent with Antweiler and Frank (2004) and Sprenger et al. (2014), messages are aligned with US market hours; specifically, messages that were posted after 4:00 pm (the market closing time) are combined, along with pre-market messages that were posted up to 9:30 am (the market opening time), with those of the following trading day since the effect of these messages on the market indicators can only appear on that day.

2.2 Classification Method of StockTwits Data

A pre-requisite step is to classify the messages into one of three distinct classes: sell, buy or hold. To manage the enormous number of the collected messages, we first choose a training data set of 2,892 tweets on all 30 DJIA stocks and classify them manually as either buy, hold or sell signals based on a pre-defined dictionary (i.e., the General Inquirer's Harvard IV-4 dictionary).[7] A few typical examples of manually classified tweets from the training data set, including the manual coding, are provided in Table 1. Next, we employ well-recognised methods from computational linguistics to accomplish the classification task based on our training data set. More specifically, unlike previous studies that use the Naive Bayes algorithm for classifying the messages (e.g., Antweiler & Frank, 2004; Sprenger et al., 2014),[8] we take a different approach by comparing the classification performances of the following three machine learning algorithms: the Naive Bayes, Decision Tree and the Support Vector Machine (SVM) algorithms.[9]

[Insert Table 1 about here]

Overall, training the selected sample of messages in Weka software using these algorithms reveals that the Random Forest Decision Tree algorithm results in a higher

that was considered by Sprenger et al. (2014).

[7] In the General Inquirer's Harvard IV-4 dictionary, more than 4,000 emotional words are tagged and classified as either positive or negative. Since a bull message indicates that an investor is optimistic and provides a "buy" signal to the market participants, it is likely that positive emotions will be associated with the "buy" class. In contrast, when an investor posts a bear message, this indicates that the investor is pessimistic and sends a "sell" signal to other market participants. The "hold" class is more likely to contain an equal balance of positive and negative emotions.

[8] The Naïve Bayes algorithm has a major drawback in sentiment classification: it assumes the conditional independence of words in documents.

[9] In normal settings, machine learning algorithms are designed to maximise the classification accuracy and minimise the error rate as much as possible (Kukar & Kononenko, 1998).

accuracy rate compared to the other algorithms. That is, applying 10-fold cross-validation (see Appendix A) reveals that the classification performance of the Decision Tree algorithm resulted in a classification accuracy of 66.7%. This is considered a good percentage, giving a random chance of 33% for the three classes (buy, sell and hold). Al-Nasseri, Menla Ali, and Tucker (2017) and its supplementary appendix provides more elaborate details about the manual labelling of these messages and their automated classification.[10]

2.3 Measurement of the Divergence of Opinion

One of the main challenges in this area of research is to find a satisfactory measure that captures differences of opinion among investors about an asset value. No measure is perfect since opinion divergence is a behavioural trait and is almost impossible to measure directly. Therefore, various proxies for the divergence of opinion among investors have been used by the extant literature. These proxies include higher trading volume (e.g., Lee & Swaminathan, 2000), breadth of mutual fund ownership (e.g., Chen et al., 2002), dispersion in analysts' earnings forecasts (e.g., Diether et al., 2002; Doukas et al., 2006; Verardo, 2009; Berkman et al., 2009; Yu, 2011; Banerjee, 2011; among others), historical income volatility, stock return volatility, firm age, average daily turnover (e.g., Berkman et al., 2009) and unexpected trading volume (e.g., Chen et al., 2015).[11]

Given the rise of microblogging platforms in exchanging information in recent years, indicators that are extracted from content that was posted on such platforms have gained popularity (see, e.g., Antweiler & Frank, 2004; Sprenger et al., 2014; Giannini et al., 2015; among others). Indeed, our adopted divergence of opinion indicator is extracted from online tweets that were posted on StockTwits. To the best of our knowledge, Sprenger et al. (2014) and Giannini et al. (2015) are the only studies to date to have used such an online indicator. However, in contrast to Sprenger et al. (2014), in which tweets that were posted on the S&P 100 stocks over a six-month period (January 1, 2010 to June 30, 2010) were classified using the Naive Bayes algorithm, we extract our indicator by using various classification algorithms to classify tweets that are related instead to the 30 DJIA highly

---

[10] Al-Nasseri et al. (2017) focuses on constructing an investor sentiment measure from this data set to examine the effect of investor sentiment on stock returns. By contrast, the focus of this paper is on investors' divergence of opinion and its impact on returns and volume.

[11] Each of these measures has its own limitations. Although dispersion in analysts' earnings forecasts is more informative, it has two potential drawbacks. First, since it only reflects the opinion divergence of professional investors, other investors tend to act independently of analysts' forecasts (i.e., their trading decisions tend to reflect their own views rather than those of the analysts). Second, analysts' forecasts are subject to the risk of uncertainty and may suffer from more behavioural biases in cases of greater information uncertainty (Zhang, 2006).

liquid stocks over a 12-month period (April 4, 2012 to April 5, 2013). Moreover, our measure differs from that used by Giannini et al. (2015) in two key ways: First, it ignores neutral (hold) and re-tweeted messages, which may introduce noise and, hence, distort the indicator. Second, in contrast to Giannini et al. (2015), who consider the level of impact that each post has in terms of the followers, which may not reflect the true or actual opinions of the platform's participants, our measure is extracted from directly posted online tweets, which reflect the true opinions of investors who are involved in the decision-making process.[12]

Having classified our messages in the previous section, following Antweiler and Frank (2004), we calculate an "agreement index", which defines the level of agreement among messages:

$$A_t = 1 - \sqrt{1 - B_t^2} \in [0, 1], \tag{1}$$

where A_t and B_t denote the agreement and bullishness indices on day *t*. The bullishness index (B_t) is an important tweet feature that determines the proportion of buy and sell signals on a particular day *t* (see Antweiler & Frank, 2004):

$$B_t = \left[\frac{M_t^{Buy} - M_t^{Sell}}{M_t^{Buy} + M_t^{Sell}}\right]. \tag{2}$$

Moreover, the agreement index is derived from the variance of B_t during time interval *t* and calculated as:

$$\sigma_t^2 = \frac{\sum_{i \in D(t)} w_i \, (x_i - B_t)^2}{\sum_{i \in D(t)} w_i} = \frac{\sum_i w_i x_i^2}{\sum_i w_i} - B_t^2 = 1 - B_t^2, \tag{3}$$

where x_i is defined as $x_i = x_i^{Buy} - x_i^{Sell} \in \{-1, +1\}$. All hold messages are excluded and ignored. w_i is the message weight. As x_i is either -1 or +1, x_i^2 will always equal 1. Thus, the simplification $\frac{\sum_i w_i x_i^2}{\sum_i w_i}$ in Eq. (3) is equal to 1. Note that when all messages are either bullish (i.e., $x_i = +1$) or bearish (i.e., $x_i = -1$) in a particular time interval, the agreement will equal 1 (since $\sqrt{1 - B_t^2}$ in Eq. (1) will equal 0). Moreover, if the number of bullish

[12] Conversations among investors in online stock forums involve making predictions, exchanging opinions, asking questions, sharing analyses, and reporting financial information (Oh & Sheng, 2011).

messages is equal to that of bearish ones in a particular time interval, the agreement will equal 0. Hence, A_t will take a value between 0 and 1. The level of agreement will decrease as the value of A_t gets closer to 0, while a high agreement level will be maintained if A_t is close to 1. Therefore, low agreement among traders' messages is interpreted as high disagreement, and vice versa. To simplify the interpretation of the results throughout, the disagreement index, which is denoted as D_t, is defined as follows:

$$D_t = -1 \times A_t. \quad (4)$$[13]

2.4 Financial Data

Daily DJIA stock closing prices and trading volumes were obtained from Bloomberg, which cover the period from April 4, 2012 to April 5, 2013. Fig. 1 displays the evolution of the DJIA index price over the sample period: No extraordinary events occurred during this period. Hence, it represents a good test base for our analysis. The DJIA is a price-weighted average of 30 large 'blue-chip' stocks that are traded on the New York Stock Exchange (NYSE) and the Nasdaq. The 30 highly traded stocks that form the index have high institutional ownership and represent approximately 25 percent of the market value of the NYSE (Lakonishok & Smidt, 1988). Furthermore, the DJIA is one of the most widely followed stock indices and attracts heavy media coverage and investor attention. For instance, Yuan (2015) documents that 92.7% of front-page market news articles refer to the DJIA index, compared to 9.4% and 0.5% for the Nasdaq and the S&P, respectively. DJIA stocks also generate a greater 'buzz' on social media networks; they are heavily discussed on the StockTwits forum and have a very high volume of tweet messages. These characteristics of such stocks make them particularly useful candidates for our empirical analysis.

[Insert Fig. 1 about here]

The return series of the DJIA stocks are computed by taking the first differences of the logarithm of the daily closing prices multiplied by 100. Since abnormal or market-adjusted returns provide a better and cleaner way to examine the relation between disagreement and stock returns beyond market directions, we also consider abnormal

[13] We thank the anonymous referee for this suggestion.

returns (AR_{it}), which is calculated as the difference between the actual returns (R_{it}) and those of the DJIA index (MKT_t):

$$AR_{it} = R_{it} - MKT_t. \tag{5}$$

Finally, the daily trading volume is the logged number of traded shares for each company on a given day *t*.

3. Empirical Methodology

We examine the impact of our constructed divergence of opinion indicator on (actual and abnormal) stock returns and trading volume using different models. The linear models that are commonly used in the literature have the following form:

$$R_{it} = \alpha_i + \lambda D_{it} + \beta MKT_t + \varepsilon_{it}, \tag{6}$$

$$AR_{it} = \alpha_i + \lambda D_{it} + \beta MKT_t + \varepsilon_{it}, \tag{7}$$

$$TV_{it} = \alpha_i + \phi TV_{it-1} + \lambda D_{it} + \beta MKT_t + \varepsilon_{it}, \tag{8}$$

where R_{it}, AR_{it} and TV_{it} are the returns, the calculated abnormal or market-adjusted returns and the trading volume, respectively, of stock i on day t; D_{it} is the level of disagreement of stock i on day t, which was calculated earlier; MKT_t is the daily return of the DJIA index, which is included as a control variable to capture the overall market-wide effects; and ε_{it} is the error term of company i on day t.[14]

Note that the panel regressions, which are shown as Eqs. (6)-(8), are estimated using the standard Ordinary Least Squares (OLS) technique and allowing for company fixed effects. Hence, they provide evidence of the linear responses of returns and volume to the divergence of opinion. Nonetheless, the effect of disagreement on returns and volume could be asymmetric, given that investor sentiment changes under different market conditions. Therefore, we also examine how such an effect differs across different states of the market:

$$R_{it} = \alpha_i + \lambda_1 I_{it}^{bull}.D_{it} + \lambda_2 I_{it}^{bear}.D_{it} + \beta MKT_t + \varepsilon_{it}, \tag{9}$$

[14] The estimated results of the volume showed that the model suffers from serial correlation. Therefore, a lag of the trading volume is also included in the model to capture its dynamics.

$$AR_{it} = \alpha_i + \lambda_1 I_{it}^{bull}.D_{it} + \lambda_2 I_{it}^{bear}.D_{it} + \beta MKT_t + \varepsilon_{it}, \tag{10}$$

$$TV_{it} = \alpha_i + \phi TV_{it-1} + \lambda_1 I_{it}^{bull}.D_{it} + \lambda_2 I_{it}^{bear}.D_{it} + \beta MKT_t + \varepsilon_{it}, \tag{11}$$

where I^{bull} and I^{bear} are indicator variables that are used to capture the impact of disagreement in bull and bear market periods, respectively, for company *i* (i.e., *i*=1,..., 30) on day *t*, which are calculated as follows:

$$I_{it}^{bull} = \begin{cases} 1 & if\ R_{it} > 0 \\ 0 & otherwise \end{cases}, \tag{12}$$

$$I_{it}^{bear} = \begin{cases} 1 & if\ R_{it} \leq 0 \\ 0 & otherwise \end{cases}. \tag{13}$$

[15]

4. Empirical Results and Discussion

In this section, we first report a summary of descriptive statistics. Then, we provide and discuss estimates of the linear and asymmetric disagreement effects on returns and volume. Finally, we provide some robustness checks of our results.

4.1. Descriptive Statistics

Table 2 reports a summary of descriptive statistics and the Pearson correlation matrix in panels A and B, respectively. The means of the actual and abnormal (or market-adjusted) stock returns are 0.045 and 0.005, respectively, while that of the trading volume (in log) is 14.59. The mean of the disagreement index is -0.116, which represents the average disagreement between the bull and bear messages on StockTwits. Note that the disagreement means during bull and bear market periods are -0.060 and -0.056, respectively. Regarding the volatility, actual stock returns are shown to have higher volatility (1.193), followed by market-adjusted returns (0.932) and trading volume (0.918), compared with disagreement (0.228).

[Insert Table 2 about here]

Furthermore, the correlations between disagreement and both actual and market-adjusted returns are shown to be insignificant and economically small. However, disagreement and trading volume show a statistically significant positive correlation (see

[15] The results (available upon request) remain unchanged, when market-adjusted returns (AR_{it}) are used instead to calculate the indicator variables.

panel B of Table 2). Thus, an increase in trading activities is associated with a higher disagreement among traders' messages. This significant positive correlation is preserved even when different market conditions are taken into consideration (i.e., bull vs. bear market periods). However, the returns and disagreement correlations during the different market states are shown to be asymmetric, i.e., significantly negative (positive) in the bull (bear) periods, which suggests that lower (higher) returns are associated with higher disagreement among traders' messages during bull (bear) periods.

Finally, Fig. 2 displays the evolution of the monthly averages of the level of disagreement, the actual stock returns, the market-adjusted stock returns and the log trading volume in panels A, B, C and D, respectively.

[Insert Fig. 2 about here]

4.2. Empirical Results

First, we report the estimates of the linear models, which are expressed as Eqs. (6)-(8) and allow for company fixed effects, in Table 3. The results show that disagreement exerts a positive but insignificant effect on both types of returns (i.e., actual and market-adjusted ones). This finding is consistent with the empirical finding in Antweiler and Frank (2004). However, it is not consistent with that of Sprenger et al. (2014) or Carlin et al. (2014), who find significant positive disagreement effect on returns, or with that of Lee and Swaminathan (2000), Diether et al. (2002), Chen et al. (2002), Berkman et al. (2009) or Yu (2011), among others, who find that the returns' reaction to disagreement is significantly negative. This limited linear effect of disagreement on returns may be due to the asymmetry of the effect. Therefore, in the next step, we further explore this effect in an asymmetric manner (i.e., over different market conditions).

[Insert Table 3 about here]

Regarding the trading volume, the results (Table 3) suggest that high disagreement among messages induces trading (e.g., λ_1= +0.066, *p*-value=0.000). In the economic sense, a 1-unit standard deviation increase in the level of disagreement (i.e., of the sell and buy messages) results in a 6.6% increase in the level of trading volume. This result supports the view of Harris and Raviv (1993) and is consistent with the empirical findings in Kandel and Pearson (1995), Antweiler and Frank (2004), Sprenger et al. (2014) and Carlin et al. (2014), among others.

Given that conclusions are commonly drawn from these linear estimates in the existing literature, further insights into the disagreement effects on returns and volume can be obtained by allowing such effects to depend on different market conditions, as in Eqs. (9)-(11). To the best of our knowledge, this is the first paper to explore such effects, even though the argument that investors behave differently in different states of the market or the economy has been well documented. Lee et al. (2002), for example, show that bullish (bearish) shifts in sentiment in the current period result in downward (upward) revisions in the volatility of future returns. More recently, Chung et al. (2012) provide evidence that, both in-sample and out-of-sample, the predictive power of investor sentiment is state-dependent; specifically, the predictive power of sentiment for the returns of portfolios formed on firm characteristics and anomalies exists in the expansion state but not in the contraction one.

We report these OLS estimates, which allow for company fixed effects, as expressed in Eqs. (9)-(11), in Table 4. It is evident that the effect of disagreement on returns is significantly asymmetric, since the effect differs across bull and bear periods. That is, disagreement exerts a negative (positive) effect on returns in the bull (bear) periods (as indicated by γ_1= -0.730, p-value=0.000 and γ_2= +0.817, p-value=0.000). This result implies that in the period of positive returns (i.e., when good news arrives), high disagreement causes lower returns; however, the effect is reversed in the period of negative returns (i.e., when bad news arrives).

The negative disagreement effect on returns in bull periods lends support to the price optimism model by Miller (1977), as it implies that divergence of opinion causes optimistic investors to drive stock prices above their intrinsic values, thereby lowering expected returns. Although Miller's (1977) intuition is grounded in the existence of short-sale constraints, which presumably do not apply to our data set, the finding is also consistent with the alternative frictionless view by Johnson (2004), who uses a general option-pricing result; specifically, for a levered firm, the expected returns should decrease with the level of idiosyncratic asset risk, which is proxied for the dispersion in analysts' earnings forecasts. In contrast, positive impact of disagreement on returns in bear periods can be explained through the concept of varying risk preference of investors (see, e.g., Varian, 1985; Kandel & Pearson, 1995; among others). For instance, unlike momentum traders, contrarians become risk lovers in the period of negative returns and are likely to be more active buyers, thereby increasing their demand of risky assets. As De Long, Shleifer, Summers, and Waldmann (1990) note, when contrarians hold more risky assets on average, they earn a larger share of returns to risk bearing; thus, their expected returns relative to those of momentum traders are higher.

[Insert Table 4 about here]

As far as the volume is concerned, the coefficients of disagreement are shown to be significantly positive in both bull and bear periods (γ_1= +0.047, *p*-value =0.06, and γ_2= +0.086, *p*-value=0.001). Specifically, a 1-unit standard deviation increase in the level of disagreement (i.e., among the buy and sell messages) leads to 4.7% and 8.6% increases in the trading volume in the bull and bear market periods, respectively. Hence, high disagreement among traders causes abnormal trading, irrespective of the state of the market. However, the magnitude of the disagreement impact is larger in bear periods compared with bull ones, which implies that disagreement among investors intensifies trading more in bear periods. This is consistent with the view that disagreement is especially intense during uncertain periods in financial markets (e.g., Hong & Stein, 2003; Carlin et al., 2014).

4.3. Robustness Checks

In this section, we carry out some robustness checks on our empirical findings by including a set of control variables in the above (linear and asymmetric) returns and volume specifications. Specifically, following Chordia, Roll, and Subrahmanyam (2001) and Antweiler and Frank (2004),[16] we include previous changes in (i) stock prices, such as stock up yesterday, stock down yesterday, stock up in the last 5 days and stock down in the last 5 days; (ii) the market index, including market up yesterday, market down yesterday, market up in the last 5 days and market down in the last 5 days; (iii) the stock and market volatility, such as stock 5-day volatility and market 5-day volatility; (iv) the trading volume, such as volume up yesterday, volume down yesterday, volume up in the last 5 days and volume down in the last 5 days; (v) the federal funds rate; (vi) the quality spread (the difference between the yield on Moody's Baa or better corporate bond and that of the 10-year Treasury bond); and (vii) the term spread (the difference between the federal funds rate and the 10-year Treasury bond yield). To capture the day-of-the-week effects, a series of dummies for Monday, Tuesday, Wednesday, and Thursday are added to the regressions. Finally, a

[16] An earlier working paper version of Sprenger et al. (2014) (referred to as Sprenger and Welpe (2010)) also follows Chordia et al. (2001) and Antweiler and Frank (2004) by including these control variables in analysing disagreement effects on volume.

dummy for each day that precedes or follows a public holiday, except when the trading day falls on a Monday or Friday, is also added.[17]

Overall, the inclusion of these control variables in the corresponding returns and volume specifications confirms our previous findings. Specifically, the results of the corresponding linear disagreement effects with the added control variables (see Table 5) suggest that the impact of disagreement on both types of returns is still insignificant and that the impact on volume remains positive and significant. Regarding the control variables, their effects on returns are broadly consistent with those observed by Chordia et al. (2001), who find that, for example, market-wide measures are more important than a company's specific measures and that interest rates play a significant role in explaining the returns (e.g., the federal funds rate, the quality spread and the term spread are all significant). The results also show a significant negative effect of the Monday dummy, thereby confirming the validity of the return anomaly of the Monday effect, which is consistent with Thaler (1987).

As for the control variable effects on volume, a firm's specific measures are shown to be relatively more prominent compared with market-wide ones, which is consistent with the findings of Antweiler and Frank (2004). However, unlike Antweiler and Frank (2004), who find elevated trading at midweek, we find all day-of-the-week dummies to be significant and negative.[18] The results also show that on Monday, there is a dramatic drop in the level of trading activities, which is indicated by the largest significant negative coefficient on the Monday dummy, compared to the other days of the week. This finding is consistent with that of Lakonishok and Maberly (1990), who observe a reduction in the trading activities of the institutional investors at the beginning of each trading week. Finally, for the interest rates, the quality spread has a significant positive effect on the trading activities.

[Insert Table 5 about here]

In contrast, the results of the corresponding asymmetric disagreement effects with the added control variables (see Table 6) suggest that disagreement exerts a significant negative (positive) influence on both types of returns in the bull (bear) periods, which is in line with our previous findings. The effects of disagreement on volume are still positive and significant in both bull and bear periods; however, their magnitudes are not significantly different across the two periods. Hence, unlike disagreement effects on returns, which are

[17] The holiday indicator takes the value of 1 for Wednesday July 4, 2012 (Independence Day), Tuesday December 25, 2012 (Christmas Day), and Thursday November 22, 2012 (Thanksgiving Day), and 0 otherwise.

[18] This finding is consistent with that in the earlier working paper version of Sprenger et al. (2014) (referred to as Sprenger and Welpe (2010)).

found to be asymmetric, these effects are symmetric. Finally, estimates of the control variables in both returns and volume specifications remain broadly unchanged in terms of sign, size and statistical significance.

[Insert Table 6 about here]

5. Returns of the Online Divergence of Opinion Exposure Portfolios

In the previous section, we examined how stock returns and trading volume react to the level of our online disagreement indicator by considering both linear and asymmetric disagreement effects. In this section, we use portfolio strategies to further verify the predictive ability of our online disagreement indicator. Specifically, we assign stocks into portfolios based on various features, such as the level of trading volume and our disagreement indicator, to explore the predictability of average returns for these groups of stocks.[19] To ensure that the results of this portfolio formation are not driven by small, illiquid stocks, or by bid-ask bounce, we follow Jegadeesh and Titman (2001), where stocks with share prices of less than five dollars are removed from the sorting process.

Following Diether et al. (2002), among others, the process is performed as follows. Each month, the stocks are assigned into five quintiles based on the volume of traded shares as of the previous month. Then, each of these portfolio groups is further sorted into five quintiles based on our online disagreement indicator as of the previous month. This two-way sorting results the stocks being assigned into 25 portfolios. The stocks are held in the portfolios for the entire trading month and are then re-sorted at the beginning of the next trading month based on new levels of volume and disagreement. The monthly portfolio returns are calculated as equally weighted average returns of all stocks in the portfolio.

The last column of Table 7 shows that if stocks are sorted by disagreement only, there is a strong positive relation between average returns and disagreement for stocks in the low to medium disagreement quintiles (i.e., when moving from D1 to D3), while a negative relation between the variables is found for stocks in the medium to high disagreement quintiles (i.e., when moving from D3 to D5); this finding is broadly in line with our previous finding of the asymmetric disagreement effect on returns. Overall, the average monthly return on the D1-D5 strategy is 0.029 percent and is significant.

[19] We additionally assign stocks into portfolios according to trading volume, because it is a well-documented predictor of returns in the literature; see, e.g., Amihud (2002), Jones (2002), and Baker and Stein (2004), among others.

The two-way sorting results that are presented in Table 7 also show that the asymmetric pattern between returns and disagreement widely prevails within each volume group. The average monthly return differential between low- and high-disagreement portfolios declines as the average trading volume increases. In particular, the returns on the D1-D5 strategy range from 0.082 for stocks in the low-volume quintile to -0.105 for stocks in the high-volume quintile, where such returns are significant for stocks in the first (low), the third (medium), and the fifth (high) volume quintiles but not for stocks in the second and fourth volume quintiles. Thus, the impact of disagreement on stock returns is greater for the relatively lower-volume companies. This result is in line with that of Sadka and Scherbina (2007), who find that the disagreement effect is particularly significant among illiquid stocks.

[Insert Table 7 about here]

In summary, portfolio strategies that are designed based on our disagreement indicator show the existence of significant profitability. The returns of portfolios of low to medium disagreement are higher than those of other portfolios, and this difference is highly significant with stocks of relatively lower trading volume.

6. Summary and Conclusions

In this paper, we shed new light on the effects of disagreement on stock market behaviour. Specifically, we analyse 289,443 online tweets from StockTwits and construct a disagreement or divergence of opinion indicator to investigate its impact on stock returns and trading volume. The data set includes the 30 DJIA highly liquid stocks over the period from April 4, 2012 to April 5, 2013. Our findings show that the stock return effects of our divergence of opinion indicator are insignificant when linear regression analysis is considered. By contrast, when we allow disagreement effects to be asymmetric, i.e., to depend on whether the market is bull (growth) or bear (decline), the returns' reaction to disagreement is significantly negative (positive) during bull (bear) periods. The existing theoretical and empirical work on the relation between disagreement and stock returns is inconclusive, with some papers predicting a negative relation (e.g., Miller, 1977; Lee & Swaminathan, 2000; Diether et al., 2002; Yu, 2011; among others) and others predicting a higher risk premium of disagreement (e.g., Varian, 1985; David, 2008; Carlin et al., 2014). To the best of our knowledge, this is the first paper to identify that a negative relation between disagreement and returns exists during bull periods and that the divergence of opinion can be viewed as a proxy for risk during bear ones.

Disagreement has also been related to higher trading volume in asset markets, both theoretically and empirically (e.g., Harris & Raviv, 1993; Kandel & Pearson, 1995; Carlin et al., 2014; among others). In this paper, we also find that this is the case and, moreover, that such a finding is supported, even when different market conditions are taken into consideration (i.e., bull vs. bear market periods).

Furthermore, portfolio strategies that are designed based on our disagreement indicator show the existence of significant profitability in the cross-section of stock returns. The returns of portfolios of low to medium disagreement are higher than those of other portfolios, and this difference is highly significant with stocks of relatively lower trading volume, which is consistent with the empirical finding of Sadka and Scherbina (2007).

Our findings have implications for related research in finance (and economics). First, they have implications for the literature on the role of social media big data in financial markets; specifically, they confirm the results of previous related studies that indicators that are extracted from such data convey valuable information that can be used in the prediction of asset prices and trading. Moreover, given that there are now various research and practical applications of the predictive ability of social media data in relation to financial indicators, future work could focus on analysing whether such data also convey information that can be utilised to predict economic indicators and other measures that are used to forecast economic activity, thereby providing important and real-time information for policy-makers.

Second, we provide consistent empirical findings for the extensive behavioural finance literature, which emphasises the roles of investors' psychology, emotions, preferences and mistaken beliefs in asset prices. Specifically, our findings further support the results of previous related studies that stock price changes and trading show sensitivity to the investor and market attention that are attracted by more visible indices, such as the DJIA (see, for example, Yuan, 2015, for recent evidence).

Third, and most importantly, our findings have implications for the specific literature on investor beliefs and their effects on asset prices. Recently, Banerjee (2011) points out that investors update their beliefs in the market based on the rational expectation approach (i.e., using information in prices efficiently) or the differences in opinion approach (i.e., agree to disagree or exhibit behavioural biases and, hence, may not condition on prices) and that these two competing approaches can be distinguished empirically by comparing how the return-volume characteristics change in relation to disagreement. Thus, our findings regarding the negative (positive) disagreement impact on returns in bull (bear) periods lend support to the differences in opinion (rational expectation) approach. In addition, the results that are obtained based on portfolio strategies confirm that investors do not update their beliefs in the market according to a single mechanism (i.e., rational expectation vs. differences in opinion). Moreover, the positive disagreement effect on volume in both bull and bear periods directly supports the view of Banerjee (2011) that higher disagreement leads to higher volume via either mechanism.

Finally, our findings have important practical implications. The explanatory power of the online disagreement indicator can be exploited by investors in their trading strategies to maximise profits and minimise the risk of their portfolios. Nonetheless, investors should keep in mind that the effects of divergence of opinion may not be the same during bull and bear periods (i.e., when different types of market news and information asymmetry among traders are at play) and for stocks with different levels of trading volume.

References

Al-Nasseri, A., Menla Ali, F., & Tucker, A. (2017). Dispersion of stock returns and investor sentiment: StockTwits evidence. Unpublished Paper.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets, 5*(1), 31-56.

Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of Internet stock message boards. *The Journal of Finance*, *59*(3), 1259-1294.

Baker, M., & Stein, J. C. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets*, *7*(3), 271-299.

Bamber, L. S., Barron, O. E., & Stober, T. L. (1997). Trading volume and different aspects of disagreement coincident with earnings announcements. *The Accounting Review, 72*(4), 575-597.

Banerjee, S. (2011). Learning from prices and the dispersion in beliefs. *The Review of Financial Studies*, *24*(9), 3025-3068.

Banerjee, S., & Kremer, I. (2010). Disagreement and learning: Dynamic patterns of trade. *The Journal of Finance, 65*(4), 1269–1302.

Berkman, H., Dimitrov, V., Jain, P. C., Koch, P. D., & Tice, S. (2009). Sell on the news: Differences of opinion, short-sales constraints, and returns around earnings announcements. *Journal of Financial Economics*, *92*(3), 376-399.

Bokányi, E., Lábszki, Z., & Vattay, G. (2017). Prediction of employment and unemployment rates from Twitter daily rhythms in the US. *EPJ Data Science*, *6*(14), Available at: https://doi.org/10.1140/epjds/s13688-017-0112-x

Bollen, J., Mao, H., & Zeng, X-J. (2011). Twitter mood predicts the stock market. *Journal Computational Science*, *2*(1), 1–8.

Bukovina, J. (2016). Social media big data and capital markets-An overview. *Journal of Behavioral and Experimental Finance*, *11*, 18–26.

Carlin, B. I., Longstaff, F. A., & Matoba, K. (2014). Disagreement and asset prices. *Journal of Financial Economics*, *114*(2), 226–238.

Chen, J., Hong, H., & Stein, J. C. (2002). Breadth of ownership and stock returns. *Journal of Financial Economics*, *66*(2-3), 171-205.

Chen, L., Qin, L., & Zhu, H. (2015). Opinion divergence, unexpected trading volume and stock returns: Evidence from China. *International Review of Economics and Finance*, *36*, 119-127.

Choi, H., & Varian, H. (2012). Predicting the present with Google trends. *Economic Record*, *88*(s1), 2–9.

Chordia, T., Roll, R., & Subrahmanyam, A. (2001). Market liquidity and trading activity. *The Journal of Finance*, *56*(2), 501-530.

Chung, S-L., Hung, C-H., & Yeh, C-Y. (2012). When does investor sentiment predict stock returns?. *Journal of Empirical Finance*, *19*(2), 217–240.

D'Amuri, F., & Marcucci, J. (2017). The predictive power of Google searches in forecasting US unemployment. *International Journal of Forecasting*, *33*(4), 801-816.

D'Avolio, G. (2002). The market for borrowing stock. *Journal of Financial Economics*, *66*(2-3), 271–306.

David, A. (2008). Heterogeneous beliefs, speculation, and the equity premium. *The Journal of Finance*, *63*(1), 41–83.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, *98*(4), 703-738.

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance*, *57*(5), 2113-2141.

Diether, K. B., Lee, K-H., & Werner, I. M. (2009). Short-sale strategies and return predictability. *The Review of Financial Studies*, *22*(2), 575–607.

Doukas, J.A., Kim, C., & Pantzalis, C. (2006). Divergence of opinion and equity returns under different states of earnings expectations. *Journal of Financial Markets*, *9*(3), 310-331.

Fang, L. H., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance*, *64*(5), 2023-2052.

Garfinkel, J. A., & Sokobin, J. (2006). Volume, opinion divergence and returns: A study of post-earnings announcement drift. *Journal of Accounting Research*, *44*(1), 85-112.

Giannini, R. C., Irvine, P. J., & Shu, T. (2015). The convergence and divergence of investors' opinions around earnings news: Evidence from a social network. Asian Finance Association (AsFA) 2013 Conference; Finance Down Under 2014 Building on the Best from the Cellars of Finance. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2017748

Harris, M., & Raviv, A. (1993). Differences of opinion make a horse race. *The Review of Financial Studies*, *6*(3), 473-506.

Hong, H., & Stein, J. C. (2003). Differences of opinion, short-sales constraints, and market crashes. *The Review of Financial Studies*, *16*(2), 487–525.

Jegadeesh, N., & Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of Finance*, *56*(2), 699-720.

Johnson, T. C. (2004). Forecast dispersion and the cross section of expected returns. *The Journal of Finance*, *59*(5), 1957-1978.

Jones, C. M. (2002). A century of stock market liquidity and trading costs. Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=313681

Kandel, E., & Pearson, N. D. (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, *103*(4), 831-872.

Kukar, M. Z., & Kononenko, I. (1998). Cost-Sensitive Learning with Neural Networks. In: Proceedings of the 13th European Conference on Artificial Intelligence (ECAI-98), pp. 445-449.

Lakonishok, J., & Maberly, E. (1990). The weekend effect: Trading patterns of individual and institutional investors. *The Journal of Finance*, *45*(1), 231-243.

Lakonishok, J., & Smidt, S. (1988). Are seasonal anomalies real? A ninety-year perspective. *The Review of Financial Studies*, *1*(4), 403-425.

Lee, C. M. C., & Swaminathan, B. (2000). Price momentum and trading volume. *The Journal of Finance*, *55*(5), 2017–2069.

Lee, W. Y., Jiang, C. X., & Indro, D. C. (2002). Stock market volatility, excess returns, and the role of investor sentiment. *Journal of Banking and Finance, 26*(12), 2277–2299.

Mao, H., Counts, S., & Bollen, J. (2011). Predicting financial markets: Comparing survey, news, twitter and search engine data. Cornell University Library. ArXiv Preprint: https://arxiv.org/abs/1112.1051

McLaren, N., & Shanbhogue, R. (2011). Using Internet search data as economic indicators. *Bank of England Quarterly Bulletin*, *51*(2), 134–140.

Milgrom, P., & Stokey, N. (1982). Information, trade and common knowledge. *Journal of Economic Theory*, *26*(1), 17-27.

Miller, E.M. (1977). Risk, uncertainty, and divergence of opinion. *The Journal of Finance*, *32*(4), 1151-1168.

Oh, C., & Sheng, O. R. L. (2011). Investigating predictive power of stock micro blog sentiment in forecasting future stock price directional movement. In *Proceedings from the 32nd International Conference on Information Systems (ICIS)*. Shanghai, China.

Sadka, R., & Scherbina, A. (2007). Analyst disagreement, mispricing, and liquidity. *The Journal of Finance*, *62*(5), 2367-2403.

Saxa, B. (2014). Forecasting mortgages: Internet search data as a proxy for mortgage credit demand. CNB Working Paper Series 14/2014.

Sprenger, T. O., & Welpe, I. M. (2010). Tweets and trades: The information content of stock microblogs. Working Paper Available at SSRN: http://ssrn.com/abstract=1702854

Sprenger, T. O., & Welpe, I. M. (2011). News or noise? The stock market reaction to different types of company-specific news events. Working Paper Available at SSRN: http://ssrn.com/abstract=1734632

Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and trades: The information content of stock microblogs. *European Financial Management*, *20*(5), 926-957.

Tetlock, P. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, *62*(3), 1139–1168.

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance, 63*(3), 1437-1467.

Thaler, R. (1987). Anomalies: Seasonal movements in security prices II: Weekend, holiday, turn of the month, and intraday effects. *The Journal of Economic Perspectives*, *1*(2), 169-177.

Varian, H. R. (1985). Divergence of opinion in complete markets: A note. *The Journal of Finance*, *40*(1), 309-317.

Verardo, M. (2009). Heterogeneous beliefs and momentum profits. *Journal of Financial and Quantitative Analysis*, *44*(4), 795-822.

Verma, R., & Verma, P. (2007). Noise trading and stock market volatility. *Journal of Multinational Financial Management*, *17*(3), 231-243.

Yu, J. (2011). Disagreement and return predictability of stock portfolios. *Journal of Financial Economics*, *99*(1), 162-183.

Yuan, Y. (2015). Market-wide attention, trading, and stock returns. *Journal of Financial Economics*, *116*(3), 548–564.

Zhang, X. F. (2006). Information uncertainty and analyst forecast behaviour. *Contemporary Accounting Research*, *23*(2), 565-590.

Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through Twitter "I hope it is not as bad as I fear". *Procedia-Social and Behavioral Sciences*, *26*, 55-62.

Zhang, W., Li, X., Shen, D., & Teglio, A. (2016). Daily happiness and stock returns: Some international evidence. *Physica A*, *460*, 201–209.

Table 1
Sample tweets from the training data set with manual classification.

Sample Tweets (Training Set)	Manual Classification
"$UNH http://stks.co/j017 Looking great for a LONG run has a position."	Buy
"$UNH $KFT http://stks.co/eANq"	Hold
"Short $NKE http://chart.ly/jmbomde"	Sell
"$KO http://stks.co/3OvK Breaks yesterday's high will add! Bullish"	Buy
"$MSFT out of the weekly 31 put"	Sell
"RT @Trade4Me Watch List: $MAT 32.01 $STT 44.74 $HON 59.16$SRZ $FSLR $DECK $CROX 22.4 $CHS $TS 36.43 $UN 33.43 $VZ $EIX $EFA $FAS $MET"	Hold
"$T good stock for buying... http://stks.co/t04i"	Buy
"$INTC $IBM Markets can use a boost from their healthy reports."	Buy
"$NKE down 3.5% wow"	Sell
"$CAT Looks ugly down there http://chart.ly/gk4hhm8"	Sell

Notes: Tweets are randomly selected and are shown in their original format. By examining the most common words that are associated with each class, we find that some general features occur frequently in all three classes (e.g., figures, ticker names and external links). However, beyond these universal features, there is a unique pattern that reasonably distinguishes the linguistic bullishness of each of the three classes. For example, positive words such as "good" and "high" are the most common words in buy messages. Financial words such as "buy", "long" and "call" give a clear indication in the financial context that the investors expected a particular stock to rise. In contrast, the most common words in sell messages are negative words such as "down", "ugly", "break" and "low", and words such as "sell", "put", "loss" and "short", which give a clear indication that the users expected the discussed stock to fall. These results match those that were observed in earlier studies (see, e.g., Tetlock, Saar-Tsechansky, & Macskassy, 2008; Sprenger et al., 2014). However, if a tweet message contains external links to long articles or charts about the stocks in which more neutral words are presented, such as the name of a product (e.g., "Aircraft", "BigMac", "Window7"), it is generally labelled hold. Therefore, in hold messages, the positive and negative words are much more equitable; that is, neutral words dominate the message.

Table 2

Summary of descriptive statistics and correlation matrix.

Panel A: Summary of descriptive statistics

	Minimum	Maximum	Mean	Std. Deviation
R_{it}	-10.96	10.49	0.045	1.193
AR_{it}	-11.11	9.867	0.005	0.932
TV_{it}	10.84	17.91	14.59	0.918
D_{it}	-1.000	0.000	-0.116	0.228
$I^{bull}.D_{it}$	-1.000	0.000	-0.060	0.175
$I^{bear}.D_{it}$	-1.000	0.000	-0.056	0.168

Panel B: Pearson correlation matrix

	R_{it}	AR_{it}	TV_{it}	D_{it}	$I^{bull}.D_{it}$	$I^{bear}.D_{it}$
R_{it}	1.000					
AR_{it}	0.780	1.000				
	(0.000)					
TV_{it}	-0.031	-0.029	1.000			
	(0.008)	(0.784)				
D_{it}	0.003	-0.002	0.048	1.000		
	(0.784)	(0.891)	(0.000)			
$I^{bull}D_{it}$	-0.225	-0.147	0.049	0.682	1.000	
	(0.000)	(0.000)	(0.000)	(0.000)		
$I^{bear}D_{it}$	0.238	0.151	0.014	0.648	-0.115	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	

Notes: The sample consists of 7,530 company-trading days for the 30 companies that make up the DJIA index over the period from April 4, 2012 to April 5, 2013. The trading volume (TV_{it}) represents the natural logarithm of the number of shares that are traded for company *i* at time *t*. Continuously compounded returns (R_{it}) are calculated and multiplied by 100. AR_{it} is the abnormal or market-adjusted stock returns. The disagreement indicator (D_{it}) is calculated and defined by Eq. (4). $I_{it}^{bull}.D_{it}$ and $I_{it}^{bear}.D_{it}$ are two interaction terms, which are used to measure the impact of disagreement during bull and bear market periods, respectively.
P-values are reported in parentheses.

Table 3

Estimates of the linear disagreement effects on returns and volume.

	R_{it}		AR_{it}		TV_{it}	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
α	0.006	(0.012)	0.006	(0.012)	10.17***	(0.160)
ϕ					0.303	(0.011)
λ	0.006	(0.048)	0.006	(0.048)	0.066***	(0.018)
β	0.998***	(0.014)	-0.002	(0.014)	-0.016***	(0.006)
R^2	0.392		0.004		0.844	
Obs	7,530		7,530		7,500	

Notes: The estimated linear effects of disagreement (D_{it}) on actual returns (R_{it}), abnormal returns (AR_{it}) and volume (TV_{it}), which allow for company fixed effects, are specified as:

$R_{it} = \alpha_i + \lambda D_{it} + \beta MKT_t + \varepsilon_{it}$,
$AR_{it} = \alpha_i + \lambda D_{it} + \beta MKT_t + \varepsilon_{it}$,
$TV_{it} = \alpha_i + \phi TV_{it-1} + \lambda D_{it} + \beta MKT_t + \varepsilon_{it}$.

MKT_t is the DJIA index return, which was added as a control variable. The sample period is April 4, 2012 to April 5, 2013. The estimated models were free from serial correlation; we added a lag in the volume specification to capture its dynamics. Standard errors are represented in parentheses.
*** denotes statistical significance at the 1% level.

Table 4
Estimates of the asymmetric disagreement effects on returns and volume.

	R_{it}		AR_{it}		TV_{it}	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
α	0.010	(0.012)	0.010	(0.012)	10.17***	(0.161)
ϕ					0.303***	(0.011)
λ_1	-0.730***	(0.062)	-0.730***	(0.062)	0.047*	(0.025)
λ_2	0.817***	(0.064)	0.817***	(0.064)	0.086***	(0.026)
β	0.932***	(0.015)	0.932***	(0.015)	-0.017***	(0.006)
R^2	0.418		0.046		0.839	
Obs	7,530		7530		7,500	

Notes: The estimated asymmetric effects of disagreement (D_{it}) on actual returns (R_{it}), abnormal returns (AR_{it}) and volume (TV_{it}), which allow for company fixed effects, are specified as:

$$R_{it} = \alpha_i + \lambda_1 I_{it}^{bull}.D_{it} + \lambda_2 I_{it}^{bear}.D_{it} + \beta MKT_t + \varepsilon_{it},$$
$$AR_{it} = \alpha_i + \lambda_1 I_{it}^{bull}.D_{it} + \lambda_2 I_{it}^{bear}.D_{it} + \beta MKT_t + \varepsilon_{it},$$
$$TV_{it} = \alpha_i + \phi TV_{i,t-1} + \lambda_1 I_{it}^{bull}.D_{it} + \lambda_2 I_{it}^{bear}.D_{it} + \beta MKT_t + \varepsilon_{it}.$$

$I_{it}^{bull}.D_{it}$ and $I_{it}^{bear}.D_{it}$ are the two interaction terms that measure the impact of disagreement during bull and bear market periods, respectively. MKT_t is the DJIA index return, which was added as a control variable. The sample period is April 4, 2012 to April 5, 2013. The estimated models were free from serial correlation; we added a lag in the volume specification to capture its dynamics. Standard errors are shown in parentheses.
* and *** denote statistical significance at the 10% and 1% levels, respectively.

Table 5
Estimates of the linear disagreement effects with additional control variables.

	R_{it}		AR_{it}		TV_{it}	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
α	-3.078***	(0.527)	0.339	(0.424)	14.45***	(0.159)
λ	0.018	(0.060)	0.016	(0.049)	0.053***	(0.018)
β_1	0.006	(0.018)	0.005	(0.015)	0.014***	(0.005)
β_2	0.018	(0.018)	0.020	(0.015)	-0.011**	(0.006)
β_3	-0.005	(0.018)	0.008	(0.015)	0.030***	(0.005)
β_4	0.008	(0.018)	-0.005	(0.015)	-0.030***	(0.005)
β_5	-0.050***	(0.018)	-0.055***	(0.015)	0.107***	(0.005)
β_6	0.155***	(0.030)	0.030	(0.025)	0.002	(0.009)
β_7	0.020	(0.031)	0.020	(0.025)	-0.042***	(0.009)
β_8	-0.030	(0.029)	-0.008	(0.024)	-0.012	(0.008)
β_9	-0.135***	(0.028)	-0.019	(0.023)	0.036***	(0.008)
β_{10}	-9.643***	(0.679)	0.096	(0.547)	1.147***	(0.205)
β_{11}	-0.161**	(0.075)	-0.068	(0.061)	0.026	(0.023)
β_{12}	0.114	(0.098)	0.015	(0.079)	0.030	(0.029)
β_{13}	0.070	(0.048)	-0.037	(0.039)	-0.066***	(0.014)
β_{14}	-0.243***	(0.057)	-0.002	(0.046)	-0.127***	(0.017)
β_{15}	1.868***	(0.407)	-0.239	(0.328)	0.209*	(0.123)
β_{16}	2.819***	(0.365)	-0.089	(0.294)	0.290***	(0.110)
β_{17}	1.431***	(0.334)	-0.268	(0.270)	0.001	(0.101)
β_{18}	-0.215***	(0.045)	0.015	(0.036)	-0.347***	(0.013)
β_{19}	0.020	(0.043)	0.006	(0.035)	-0.273***	(0.013)
β_{20}	-0.094**	(0.042)	0.030	(0.034)	-0.260***	(0.012)
β_{21}	-0.064	(0.042)	-0.006	(0.034)	-0.191***	(0.012)
β_{22}	0.126	(0.126)	-0.034	(0.102)	-0.437***	(0.038)
R^2	0.055		0.008		0.855	
Obs	7,410		7,410		7,410	

Notes: The estimated linear effects of disagreement on (actual and market-adjusted) returns and volume, which allow for company fixed effects and include a set of additional control variables, are specified as:

$$R_{it} = \alpha_i + \lambda D_{it} + \beta_1 Stock_{it}{}^{+} + \beta_2 Stock_{it}{}^{-} + \beta_3 MA5\ Stock_{it}{}^{+} + \beta_4 MA5\ Stock_{it}{}^{-} + \beta_5 MA5\ stock\ volt_{it} + \beta_6 MKT_t{}^{+} + \beta_7 MKT_t{}^{-} + \beta_8 MA5\ MKT_t{}^{+} + \beta_9 MA5\ MKT_t{}^{-} + \beta_{10} MA5\ MKT\ volt_t + \beta_{11} TV_t{}^{+} + \beta_{12} TV_t{}^{-} + \beta_{13} MA5\ TV_{it}{}^{+} + \beta_{14} MA5\ TV_{it}{}^{-} + \beta_{15}\Delta FFR_t + \beta_{16} Qlty\ Sprd_t + \beta_{17} Trm\ Sprd_t + \beta_{18} MON_t + \beta_{19} TUE_t + \beta_{20} WED_t + \beta_{21} THU_t + \beta_{22} Holiday_t + \varepsilon_{it},$$

$$AR_{it} = \alpha_i + \lambda D_{it} + \beta_1 Stock_{it}{}^{+} + \beta_2 Stock_{it}{}^{-} + \beta_3 MA5\ Stock_{it}{}^{+} + \beta_4 MA5\ Stock_{it}{}^{-} + \beta_5 MA5\ stock\ volt_{it} + \beta_6 MKT_t{}^{+} + \beta_7 MKT_t{}^{-} + \beta_8 MA5\ MKT_t{}^{+} + \beta_9 MA5\ MKT_t{}^{-} + \beta_{10} MA5\ MKT\ volt_t + \beta_{11} TV_t{}^{+} + \beta_{12} TV_t{}^{-} + \beta_{13} MA5\ TV_{it}{}^{+} + \beta_{14} MA5\ TV_{it}{}^{-} + \beta_{15}\Delta FFR_t + \beta_{16} Qlty\ Sprd_t + \beta_{17} Trm\ Sprd_t + \beta_{18} MON_t + \beta_{19} TUE_t + \beta_{20} WED_t + \beta_{21} THU_t + \beta_{22} Holiday_t + \varepsilon_{it},$$

$$TV_{it} = \alpha_i + \lambda D_{it} + \beta_1 Stock_{it}{}^{+} + \beta_2 Stock_{it}{}^{-} + \beta_3 MA5\ Stock_{it}{}^{+} + \beta_4 MA5\ Stock_{it}{}^{-} + \beta_5 MA5\ stock\ volt_{it} + \beta_6 MKT_t{}^{+} + \beta_7 MKT_t{}^{-} + \beta_8 MA5\ MKT_t{}^{+} + \beta_9 MA5\ MKT_t{}^{-} + \beta_{10} MA5\ MKT\ volt_t + \beta_{11} TV_t{}^{+} + \beta_{12} TV_t{}^{-} + \beta_{13} MA5\ TV_{it}{}^{+} + \beta_{14} MA5\ TV_{it}{}^{-} + \beta_{15}\Delta FFR_t + \beta_{16} Qlty\ Sprd_t + \beta_{17} Trm\ Sprd_t + \beta_{18} MON_t + \beta_{19} TUE_t + \beta_{20} WED_t + \beta_{21} THU_t + \beta_{22} Holiday_t + \varepsilon_{it}.$$

R_{it}, AR_{it} and TV_{it} are the actual returns, abnormal returns and trading volume, respectively, for the ith stock on day t. D_{it} is the disagreement indicator. If we let $P_t^c(P_t^o)$ and $\bar{P}_t^c(\bar{P}_t^o)$ denote, respectively, the closing (opening) price of a particular stock and the closing (opening) price of the market index, the control variables are referred to and calculated as follows: $Stock_t{}^{+}$ is stock up yesterday = max{0, ln(P_{t-1}^c)- ln(P_{t-2}^c)}; $Stock_t{}^{-}$ is stock down yesterday = max{0, ln(P_{t-2}^c)- ln(P_{t-1}^c)}; $MA5\ Stock_t{}^{+}$ is stock up in the last 5 days = max{0, ln(P_{t-1}^c) - ln(P_{t-5}^o) }; $MA5\ Stock_t{}^{-}$ is stock down in the last 5 days = max{0, ln(P_{t-5}^o) - ln(P_{t-1}^c) }; $MA5\ stock\ volt_t$ is stock five-day volatility = $\sum_{q=1}^{5}\sum_{d\,\in\,D(t)}\left|\ln\left(P_{t-q,d}\right) - \ln(P_{t-q,d-1})\right|$; $MKT_t{}^{+}$is market up yesterday = max{0, ln($\bar{P}_{t-1}^c$)- ln($\bar{P}_{t-2}^c$)}; $MKT_t{}^{-}$ is market down yesterday = max{0, ln($\bar{P}_{t-2}^c$)- ln($\bar{P}_{t-1}^c$)}; $MA5\ MKT_t{}^{+}$ is market up in the last 5 days = max{0, ln($\bar{P}_{t-1}^c$)- ln($\bar{P}_{t-5}^o$)}; $MA5\ MKT_t{}^{-}$ is market down in the last 5 days = max{0, ln($\bar{P}_{t-5}^o$) - ln($\bar{P}_{t-1}^c$) }; $MA5\ MKT\ volt_t$ is market five-day volatility =

$\sum_{i=1}^{5}\sum_{d\,\in\,D(t)}\left|\ln\left(\bar{P}_{t-i,d}\right)-\ln(\bar{P}_{t-i,d-1})\right|$; ${TV_t}^{+}$ is volume up yesterday = max$\{0$, ln(TV_{t-1})- ln$(TV_{t-2})\}$; ${TV_t}^{-}$ is volume down yesterday = max$\{0$, ln(TV_{t-2})- ln$(TV_{t-1})\}$; $MA5\,{TV_t}^{+}$ is volume up in the last 5 days = max$\{0$, ln(TV_{t-1})- ln$(TV_{t-5})\}$; $MA5\,{TV_t}^{-}$ is volume down in the last 5 days = max$\{0$, ln(TV_{t-5})- ln$(TV_{t-1})\}$; ΔFFR_t is the change in the federal funds rate (i.e., ln(FFR_t)- ln(FFR_{t-1})); $Qlty\ Sprd_t$ is the quality spread, which is calculated as the change in the yield differential between the BAA corporate bond (BAA_t) and the 10-year Treasury bond ($T10_t$) (i.e., $\Delta(BAA_t - T10_t)$; $Trm\ Sprd_t$ is the term spread, which is calculated as the change in the interest rate differential between the 10-year Treasury bond and the federal funds rate (i.e., $\Delta(T10_t - FFR_t)$; *MON, TUE, WED, THU* are dummies for different days of the week, which take the value of 1 if the trading day is Monday, Tuesday, Wednesday or Thursday, and 0 otherwise; and *Holiday* is a dummy variable, which takes the value of 1 if the trading day is Wednesday July 4, 2012 (Independence Day), Tuesday December 25, 2012 (Christmas Day), or Thursday November 22, 2012 (Thanksgiving Day), and 0 otherwise. The sample period is April 4, 2012 to April 5, 2013. The estimated models were free from serial correlation. Standard errors are presented in parentheses.

*, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6

Estimates of the asymmetric disagreement effects with additional control variables.

	R_{it}		AR_{it}		TV_{it}	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
α	-2.806***	(0.503)	0.488	(0.416)	14.45***	(0.159)
λ_1	-1.246***	(0.074)	-0.676***	(0.062)	0.052**	(0.023)
λ_2	1.492***	(0.079)	0.822***	(0.066)	0.054**	(0.025)
β_1	0.006	(0.018)	0.005	(0.015)	0.014***	(0.005)
β_2	0.013	(0.018)	0.017	(0.015)	-0.011**	(0.005)
β_3	-0.009	(0.017)	0.006	(0.015)	0.030***	(0.005)
β_4	0.010	(0.017)	-0.003	(0.015)	-0.030***	(0.005)
β_5	-0.048***	(0.017)	-0.055***	(0.015)	0.107***	(0.005)
β_6	0.128***	(0.029)	0.015	(0.024)	0.002	(0.009)
β_7	0.023	(0.029)	0.021	(0.025)	-0.043***	(0.009)
β_8	-0.021	(0.028)	-0.003	(0.023)	-0.012	(0.008)
β_9	-0.128***	(0.027)	-0.015	(0.023)	0.037***	(0.008)
β_{10}	-8.804***	(0.649)	0.555	(0.524)	1.148***	(0.205)
β_{11}	-0.134*	(0.072)	-0.053	(0.060)	0.026	(0.023)
β_{12}	0.074	(0.093)	-0.007	(0.077)	0.030	(0.029)
β_{13}	0.050	(0.046)	-0.048	(0.038)	-0.066***	(0.014)
β_{14}	-0.195***	(0.055)	0.025	(0.045)	-0.128***	(0.017)
β_{15}	1.541***	(0.388)	-0.418	(0.321)	0.210*	(0.123)
β_{16}	2.599***	(0.349)	-0.209	(0.288)	0.290**	(0.110)
β_{17}	1.294***	(0.319)	-0.343	(0.264)	0.001	(0.101)
β_{18}	-0.170***	(0.043)	0.040	(0.036)	-0.347***	(0.013)
β_{19}	0.023	(0.041)	0.009	(0.035)	-0.273***	(0.013)
β_{20}	-0.088**	(0.040)	0.033	(0.034)	-0.261***	(0.012)
β_{21}	-0.057	(0.040)	-0.002	(0.034)	-0.191***	(0.012)
β_{22}	0.084	(0.120)	-0.056	(0.099)	-0.437***	(0.038)
R^2	0.138		0.048		0.855	
Obs	7,410		7, 410		7,410	

Notes: The estimated asymmetric effects of disagreement on (actual and market-adjusted) returns and volume allowing for company fixed effects and including a set of additional control variables are specified as:

$$R_{it} = \alpha_i + \lambda_1 I_{it}^{bull}.D_{it} + \lambda_2 I_{it}^{bear}.D_{it} + \beta_1 Stock_{it}^{+} + \beta_2 Stock_{it}^{-} + \beta_3 MA5\ Stock_{it}^{+} + \beta_4 MA5\ Stock_{it}^{-} + \beta_5 MA5\ stock\ volt_{it} + \beta_6 MKT_t^{+} + \beta_7 MKT_t^{-} + \beta_8 MA5\ MKT_t^{+} + \beta_9 MA5\ MKT_t^{-} + \beta_{10} MA5\ MKT\ volt_t + \beta_{11} TV_t^{+} + \beta_{12} TV_t^{-} + \beta_{13} MA5\ TV_{it}^{+} + \beta_{14} MA5\ TV_{it}^{-} + \beta_{15} \Delta FFR_t + \beta_{16} Qlty\ Sprd_t + \beta_{17} Trm\ Sprd_t + \beta_{18} MON_t + \beta_{19} TUE_t + \beta_{20} WED_t + \beta_{21} THU_t + \beta_{22} Holiday_t + \varepsilon_{it},$$

$$AR_{it} = \alpha_i + \lambda_1 I_{it}^{bull}.D_{it} + \lambda_2 I_{it}^{bear}.D_{it} + \beta_1 Stock_{it}^{+} + \beta_2 Stock_{it}^{-} + \beta_3 MA5\ Stock_{it}^{+} + \beta_4 MA5\ Stock_{it}^{-} + \beta_5 MA5\ stock\ volt_{it} + \beta_6 MKT_t^{+} + \beta_7 MKT_t^{-} + \beta_8 MA5\ MKT_t^{+} + \beta_9 MA5\ MKT_t^{-} + \beta_{10} MA5\ MKT\ volt_t + \beta_{11} TV_t^{+} + \beta_{12} TV_t^{-} + \beta_{13} MA5\ TV_{it}^{+} + \beta_{14} MA5\ TV_{it}^{-} + \beta_{15} \Delta FFR_t + \beta_{16} Qlty\ Sprd_t + \beta_{17} Trm\ Sprd_t + \beta_{18} MON_t + \beta_{19} TUE_t + \beta_{20} WED_t + \beta_{21} THU_t + \beta_{22} Holiday_t + \varepsilon_{it},$$

$$TV_{it} = \alpha_i + \lambda_1 I_{it}^{bull}.D_{it} + \lambda_2 I_{it}^{bear}.D_{it} + \beta_1 Stock_{it}^{+} + \beta_2 Stock_{it}^{-} + \beta_3 MA5\ Stock_{it}^{+} + \beta_4 MA5\ Stock_{it}^{-} + \beta_5 MA5\ stock\ volt_{it} + \beta_6 MKT_t^{+} + \beta_7 MKT_t^{-} + \beta_8 MA5\ MKT_t^{+} + \beta_9 MA5\ MKT_t^{-} + \beta_{10} MA5\ MKT\ volt_t + \beta_{11} TV_t^{+} + \beta_{12} TV_t^{-} + \beta_{13} MA5\ TV_{it}^{+} + \beta_{14} MA5\ TV_{it}^{-} + \beta_{15} \Delta FFR_t + \beta_{16} Qlty\ Sprd_t + \beta_{17} Trm\ Sprd_t + \beta_{18} MON_t + \beta_{19} TUE_t + \beta_{20} WED_t + \beta_{21} THU_t + \beta_{22} Holiday_t + \varepsilon_{it}.$$

$I_{it}^{bull}.D_{it}$ and $I_{it}^{bear}.D_{it}$ are the two interaction terms that measure the impact of disagreement during bull and bear market periods, respectively. The notations of the rest of the variables are defined in the notes of Table 5. The sample period is April 4, 2012 to April 5, 2013. The estimated models were free from serial correlation. Standard errors are presented in parentheses.
*, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 7

Mean portfolio returns by volume and disagreement.

		Mean Returns					
		Volume Quintiles					
		V1 (small)	V2	V3	V4	V5 (large)	All Stocks
Disagreement Quintiles	D1 (low)	0.105	0.082	-0.090	0.138	-0.097	0.026
	D2	0.169	0.027	0.038	0.074	-0.042	0.055
	D3	0.169	0.119	0.018	0.208	0.024	0.107
	D4	0.057	0.131	0.015	0.034	-0.047	0.032
	D5 (high)	0.023	0.075	-0.004	0.187	0.008	0.055
	D1 – D5	0.082** (3.124)	0.007 (0.427)	-0.086*** (-4.299)	-0.049 (-1.662)	-0.105*** (-5.450)	0.029** (-2.216)

Notes: The two-way portfolio sorting is based on trading volume and the level of our online disagreement indicator. Each month, stocks are sorted into five quintiles based on the trading volume at the end of the previous month. Stocks in each volume quintile are then re-sorted into an additional five quintiles based on the level of disagreement in the previous month. Hence, stocks are assigned into 25 portfolio groups, which are then held for one month. The monthly weighted average returns are calculated for each portfolio group. The sample period is April 4, 2012 to April 5, 2013. V1 (V5) is comprised of the stocks with the lowest (highest) volume of trade, while D1 (D5) contains the stocks with low (high) disagreement. The t-statistics, which are presented in parentheses, are based on a mean difference of zero.
** and *** indicate statistical significance at the 5% and 1% levels, respectively.

Fig. 1. Evolution of the daily closing prices of the DJIA index over the period from April 4, 2012 to April 5, 2013.

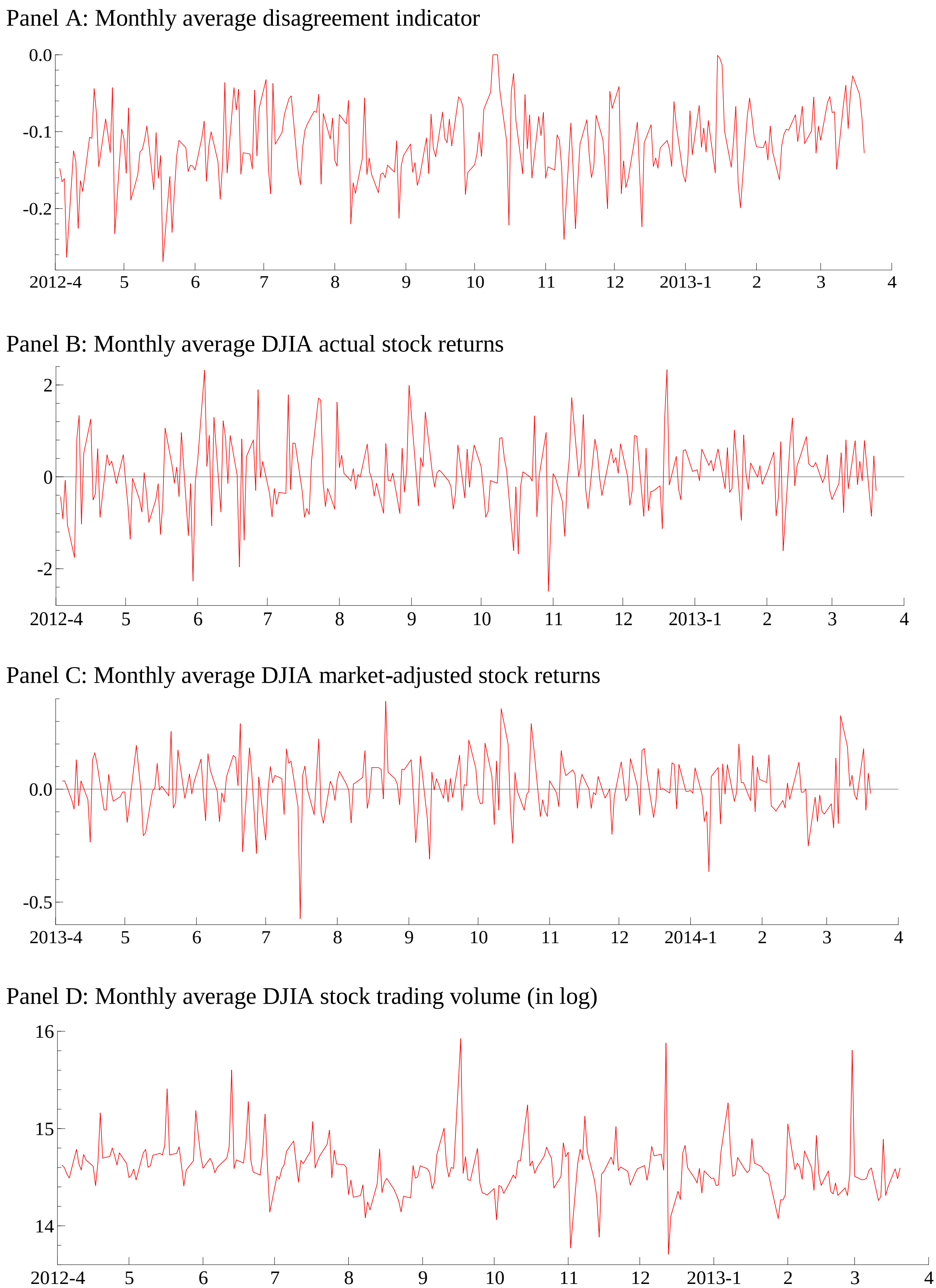

Fig. 2. Monthly averages of the level of disagreement (Panel A), the actual returns (Panel B), the market-adjusted returns (Panel C) and the log trading volume (Panel D) over the period from April 2012 to April 2013.

Appendix A: Performance Analysis of the Random Forest Decision Tree Classifier

Accuracy Rate

The Random Forest Decision Tree classifier was tested using the full training data set (with 2,892 tweets) and 10-fold cross-validation. The results (see Table A1) show that testing on the training data set produced an accuracy rate of 97.72% (i.e., 2,826 instances were correctly classified out of 2,892), whereas testing using 10-fold cross-validation generated an accuracy rate of 66.70% (i.e., 1,929 instances were correctly classified out of 2,892). Overall, this indicates that the constructed Random Forest model has learnt the training data set extremely well (given the high accuracy rate of 97.72%). Moreover, since there were three classes in the target variables (buy, sell and hold), any percentage that is greater than 33% (i.e., the probability of occurrence of each class is 1/3= 0.33= 33%) is considered good. Hence, the 66.70% accuracy rate by 10-fold cross-validation is well within the desired range and gives a random-chance probability of 33% for each class.

[Insert Table A1 about here]

Classification Accuracy by Class

Table A2 reports the classification accuracy by class for the Random Forest Decision Tree classifier, including the cost-insensitive measures (e.g., precision, recall and F-measure) and the cost-sensitive ones (e.g., the receiver operating characteristic (ROC) area[20]) for each class. Testing directly on the full training data set yielded an accuracy rate of 97.72% (see Table A1). However, such testing will always show better results compared with what is expected from stratified cross-validation with 10 folds (which is considered a more conservative measure of the classification accuracy). This is because k-fold cross-validation provides the best generalisation ability and helps overcome the risk of model over-fitting, as each of these folds checks if the learnt model is over-fitted on the validation set. Accordingly, stratified cross-validation provides a more realistic picture than testing on the full training data set. Therefore, our main focus in the analysis is on the results of 10-fold cross-validation, to strengthen the validity of our findings. Consistent with standard metrics of information retrieval, measures such as recall, precision and F-measure are reported and

[20] The ROC graph is a technique that is used to analyse, compare, organise and select classifiers based on their performances (Fawcett, 2006; Prati, Batista, & Monard, 2004). The ROC is used to perform comparative analysis by evaluating the sensitivities and specificities of different classifiers. An ROC graph shows a trade-off between the sensitivity (hit rate) and specificity (false-alarm rate) of each classifier (Swets, Dawes, & Monahan, 2000). It is used to measure the area under the curve (AUC).

used to evaluate the performance of the predictive model (Witten & Frank, 2005).

[Insert Table A2 about here]

As shown in Table A2, when using stratified cross-validation with 10 folds, the buy class has a precision of 69.60%, a recall of 78.40%, and an F-measure of 73.70%, which indicates very good performance by the Random Forest model in predicting this class. The hold class has a precision of 61.60%, a recall of 63.90%, and an F- measure of 62.70%. The sell class has a precision of 65.00%, but the recall is relatively low (51.50%), which causes the F-measure to drop to 57.50%. The weighted averages of the three classes show comparable results (65.50%, 66.70% and 66.20% for precision, recall and F-measures, respectively). Moreover, the areas under the curve (AUCs) for Random Forest Tree are 80.00%, 85.90% and 75.70% for the buy, hold and sell classes, respectively. Note that the closer the AUC of a class is to 100%, the better the classifier is in predicting that class. Therefore, Random Forest Tree achieves relatively good classification accuracy, as indicated by the AUC (i.e., scored > 75.00% in all three classes). It is worth noting that the AUC for the hold class is higher than those of the buy and sell classes, which suggests higher accuracy of the hold-class classification compared to the other classes.

Confusion Matrix

The confusion matrix is one of the most important performance evaluation measures of machine learning models. According to Table A3, when using the full training data set, the Random Forest model shows excellent classification accuracy for all three classes. The prediction accuracies for the buy, hold and sell classes are 99.50% (1,354/1361), 97.10% (573/590) and 95.50% (899/941), respectively. Note that the numbers of manually classified tweets are 1,361, 590, and 941 for the buy, hold and sell classes, respectively.

[Insert Table A3 about here]

The confusion matrix for the 10-fold cross-validation that was presented in the second part of Table A3 shows the classification of the instances (messages) from each class in the training data set. For example, the buy class has 1,361 buy instances (i.e., 1,361 = 1,067+117+177) in our training data set, of which the classification model correctly classified 1,067 as buy but incorrectly identified 294 instances (i.e., 294 = 117+177) as hold or sell. Therefore, 1,067 instances are true positives and 294 instances are false negatives of the buy class. For the hold class, we have 590 hold instances in the training sample, of which the model correctly classified 377 instances and misclassified 213 instances (i.e., 213

= 129+84; 129 and 84 instances were incorrectly assigned to the buy and sell classes, respectively). Hence, 377 and 213 instances are true positives and false negatives, respectively. Finally, for the sell class, the model correctly classified 485 instances (i.e., true-positive instances) and misclassified 338 and 118 instances (i.e., false-negative instances) to the buy and hold classes, respectively. Overall, the Random Forest Decision Tree model shows the highest classification accuracy for messages that belong to the buy class, but it was less successful at predicting the messages in the sell class.

Appendix A References

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861-874.

Prati, R. C., Batista, G. E., & Monard, M. C. (2004). Class imbalances versus class overlapping: An analysis of a learning system behaviour. In: R. Monroy, G. Arroyo-Figueroa, L.E. Sucar, & H. Sossa (Eds.), *MICAI 2004: Advances in Artificial Intelligence* (pp. 312-321). Mexico: Springer.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Better decisions through science. *Scientific American, 283*(4), 82-87.

Witten, I. H., & Frank, E. (2005). *Data Mining: Practical Machine Learning Tools and Techniques* (2nd ed). San Francisco, CA: Morgan Kaufmann.

Table A1
Weka summary results for the Random Forest Decision Tree classifier.

Testing Method	Accuracy Rate	Correctly Classified Instances	Incorrectly Classified Instances
Full Training Data Set	97.72%	2,826	66
10-Fold Cross-Validation	66.70%	1,929	963

Table A2
Classification accuracy by class using the Random Forest Decision Tree classifier.

Panel A: Full Training Data Set

Class	True Positives	False Positives	Precision	Recall	F-Measure	ROC Area
Buy	99.50%	2.90%	96.80%	99.50%	98.10%	99.90%
Hold	97.10%	0.40%	98.30%	97.10%	97.70%	99.90%
Sell	95.50%	0.60%	98.80%	95.50%	97.10%	99.80%
Weighted Average	97.70%	1.70%	97.70%	97.70%	97.70%	99.90%

Panel B: 10-Fold Cross-Validation

Class	True Positives	False Positives	Precision	Recall	F-Measure	ROC Area
Buy	78.40%	30.50%	69.60%	78.40%	73.70%	80.00%
Hold	63.90%	10.20%	61.60%	63.90%	62.70%	85.90%
Sell	51.50%	13.40%	65.00%	51.50%	57.50%	75.70%
Weighted Average	66.70%	20.80%	66.50%	66.70%	66.20%	79.80%

Notes: True positives represent the messages that were correctly classified to a given class. False positives are messages that were classified incorrectly to a given class. Precision is the proportion of the messages that were correctly classified to a class of all messages that were classified to that class. The recall (also known as sensitivity) of a class represents the share of all messages that were classified correctly. Note that the Recall measure is equivalent to the True-Positive rate. F-measure is a combined measure for precision and recall and is calculated as F= (2* Precision* Recall) / (Precision +Recall). The ROC Area measure is one of the most important measures of Weka output. It exemplifies the performance of the classification model through the trade-off between the classifier sensitivity (TP_{rate}) and specificity (false alarm rate FP_{rate}), where the sensitivity can be increased with a small loss in specificity, and vice versa.

Table A3
Classification accuracy (confusion matrix) for the Random Forest Decision Tree classifier.

Panel A: Full Training Data Set

Classified As	Buy	Hold	Sell
Buy	1,354	2	5
Hold	11	573	6
Sell	34	8	899

Panel B: 10-Fold Cross-Validation

Classified As	Buy	Hold	Sell
Buy	1,067	117	177
Hold	129	377	84
Sell	338	118	485

Notes: Each element in the matrix is a count of instances. The rows represent the correctly classified instances (messages) of a given class, while the columns represent the predicted instances of that class.